Theor Appl Genet (2014) 127:1365–1374
DOI 10.1007/s00122-014-2304-0

ORIGINAL PAPER

A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds

Shin Kato · Takashi Sayama · Kenichiro Fujii · Setsuzo Yumoto · Yuhi Kono · Tae-Young Hwang · Akio Kikuchi · Yoshitake Takada · Yu Tanaka · Tatsuhiko Shiraiwa · Masao Ishimoto

Received: 19 November 2013 / Accepted: 24 March 2014 / Published online: 10 April 2014

Abstract

Key message **We detected a QTL for single seed weight in soybean that was stable across multiple environments and genetic backgrounds with the use of two recombinant inbred line populations.**

Abstract Single seed weight (SSW) in soybean is a key determinant of both seed yield and the quality of soy food products, and it exhibits wide variation. SSW is under genetic control, but the molecular mechanisms of such control remain unclear. We have now investigated quantitative trait loci (QTLs) for SSW in soybean and have identified such a QTL that is stable across multiple environments and genetic backgrounds. Two populations of 225 and 250 recombinant inbred lines were developed from crosses between Japanese and US cultivars of soybean that differ in SSW by a factor of ~2, and these populations were grown in at least three different environments. A whole-genome panel comprising 304 simple sequence repeat (SSR) loci was applied to mapping in each population. We identified 15 significant QTLs for SSW dispersed among 11 chromosomes in the two populations. One QTL located between Sat_284 and Sat_292 on chromosome 17 was detected ($3.6 < \text{LOD} < 14.1$) in both populations grown in all environments. This QTL, tentatively designated *qSw17-1*, accounted for 9.4–20.9 % of phenotypic variation in SSW, with a dominant allele being associated with increased SSW. Given its substantial effect on SSW, *qSw17-1* is an attractive target for positional cloning, and SSR markers closely associated with this locus may prove useful for marker-assisted selection for SSW control in soybean.

Communicated by B. Diers.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2304-0) contains supplementary material, which is available to authorized users.

S. Kato · S. Yumoto · Y. Kono · A. Kikuchi · Y. Takada
National Agriculture and Food Research Organization (NARO) Tohoku Region Agricultural Research Center, 297 Uenodai, Kariwano, Daisen, Akita 019-2112, Japan

T. Sayama · M. Ishimoto (✉)
National Institute of Agrobiological Sciences (NIAS), 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan
e-mail: ishimoto@affrc.go.jp

T. Sayama · T.-Y. Hwang · M. Ishimoto
NARO Hokkaido Region Agricultural Research Center, 1 Hitsujigaoka, Toyohira, Sapporo, Hokkaido 062-8555, Japan

K. Fujii · Y. Tanaka · T. Shiraiwa
Graduate School of Agriculture, Kyoto University, Oiwake, Kitashirakawa, Sakyo, Kyoto 606-8502, Japan

Present Address:
Y. Kono
NARO Kyushu Okinawa Agricultural Research Center, 2421 Suya, Koshi, Kumamoto 861-1192, Japan

Present Address:
T.-Y. Hwang
National Institute of Animal Science, RDA, Cheonan 330-801, Republic of Korea

Present Address:
Y. Takada
NARO Western Region Agricultural Research Center, 1-3-1 Senyu, Zentsuji, Kagawa 765-8508, Japan

Introduction

Single seed weight (SSW), often expressed as 100-seed weight, is an important yield component in soybean and has been found to show a positive correlation with seed yield (Burris et al. 1973; Smith and Camper 1975). SSW is also associated with germination viability and seed vigor (Edwards and Hartwig 1971; Hopper et al. 1979). In addition, seed size, or SSW, is a determinant of the quality of soy food products, with small-seed soybean generally being used for natto and soy sprouts and large-seed soybean being preferred for tofu, edamame, and boiled soybean. As a result, soybean germplasm manifests wide variation in SSW, which ranges 30–775 mg per seed (Hartwig 1973; Kaga et al. 2012).

SSW in soybean is a transmittable trait with high heritability ranging 44–94 % (Brim 1973), and it is controlled by many genes with additive effects (Brim and Cockerham 1961). The construction of consensus linkage maps for soybean (Song et al. 2004; Hwang et al. 2009; Hyten et al. 2010) has facilitated the identification of agronomic trait loci, including quantitative trait loci (QTLs). This strategy has been applied to the detection of QTLs for SSW, with several such loci having been identified (Orf et al. 1999; Chapman et al. 2003; Hoeck et al. 2003; Fasoula et al. 2004; Hyten et al. 2004; Zhang et al. 2004; Panthee et al. 2005; Xu et al. 2011; Han et al. 2012; Hu et al. 2013). A recent genome-wide association study performed with 257 soybean cultivars also described the detection of QTLs associated with seed size (Niu et al. 2013). Indeed, many QTLs for SSW that map to all the 20 chromosomes have been deposited to date in SoyBase (http://soybase.org, accessed 27 October 2013). These findings indicate that SSW in soybean is controlled by a series of QTLs that function in different combinations and strengths in different populations or environments. However, with the exception of one locus, *ln* (narrow leaf) (Mandl and Buss 1981; Dinkins et al. 2002; Jeong et al. 2012), no gene that controls SSW has been isolated.

Recent analysis of allelic diversity at molecular marker loci has revealed that the Japanese and Chinese cultivars of soybean constitute different germplasm pools (Zhou et al. 2000; Hwang et al. 2008; Kaga et al. 2012). On the other hand, North American as well as southeast and south-central Asian cultivars overlap with the diverse Chinese germplasm pool (Abe et al. 2003; Ude et al. 2003). The Japanese germplasm pool likely reflects relatively few introductions from the ancestral continental germplasm pool, with subsequent breeding efforts having led to the development of unique and distinct germplasm pools in North America and Japan. Given that SSW is a commercially important trait for Japanese traditional soy food products such as natto, edamame, and boiled soybean, ongoing and sustained selection might have resulted in the wide variation in SSW apparent in the Japanese germplasm pool in spite of its relatively low genetic diversity (Kaga et al. 2012). Unique genes that control SSW may remain or have emerged in the Japanese germplasm pool. However, genetic analysis of SSW performed with Japanese cultivars has been limited (Watanabe et al. 2004).

The objective of the present study was to detect QTLs for SSW in soybean that are substantial and stable across multiple environments and genetic backgrounds. For this purpose, we developed two recombinant inbred line (RIL) populations from crosses between US standard soybean cultivars (Athow and Stressland) and typical Japanese large-seed soybean cultivars (Ohsuzu and Tachinagaha), which differ by a factor of ~2 in SSW, and we grew these populations in at least three different environments.

Materials and methods

Plant materials

Two RIL populations, designated OA-RILs and ST-RILs, were developed by the single seed descent (SSD) method from Ohsuzu × Athow (PI 595926) and Stressland (PI 593654) × Tachinagaha (PI 561396) crosses, respectively. Athow [maturity group (MG) III] and Stressland (MG IV) were selected as standard US cultivars developed for soybean meal and vegetable oil by the Agricultural Research Service of the US Department of Agriculture (USDA-ARS) (Wilcox and Abney 1997; Cooper et al. 1999). Ohsuzu (MG III) and Tachinagaha (MG V) were selected as typical Japanese large-seed cultivars used for food products such as tofu and boiled (cooked) beans. Generations were advanced to F_6 for the OA-RIL population and to F_7 for the ST-RIL population by SSD, after which they were advanced without selection. The OA-RIL population consists of 225 lines, and the ST-RIL population comprises 250 lines. To determine the mode of inheritance of a main QTL, we selected heterozygous F_6 plants from the OA-RIL population using DNA markers in the region proximal to the QTL. These plants were selfed to produce $F_{6:7}$ residual heterozygous lines (RHLs) (Yamanaka et al. 2005). Segregating seeds were obtained by selfing of a heterozygous individual in each selected line in 2007.

Experimental conditions of field trials

All plant materials with the exception of OA-RHLs were grown in two locations, Akita (the Kariwano branch of the Daisen research station of NARO Tohoku Region Agricultural Research Center, located at 39°32′N, 140°22′E) and Kyoto (the Experimental Farm of Kyoto University, located

Table 1 Statistical analysis of single seed weight (SSW) in the parental cultivars and the OA-RIL and ST-RIL populations grown in different environments

Population	Environment (location-year)	SSW (mg/seed)						
		Parents		RILs				
		P1	P2	Range	Mean	CV	W	
OA-RIL								
P1: Ohsuzu	Akita-2007	373	200	205–403	276	13.1	0.99	NS
P2: Athow	Akita-2008	302	177	165–338	234	13.1	0.99	NS
	Akita-2010	341	211	197–327	250	10.9	0.99	NS
	Kyoto-2010	292	179	174–328	246	12.1	0.99	NS
ST-RIL								
P1: Stressland	Akita-2010	153	352	153–305	236	13.5	0.98	NS
P2: Tachinagaha	Kyoto-2010	166	358	180–323	241	14.3	0.96	**
	Akita-2011	170	358	146–317	230	14.6	1.00	NS

W indicates statistical value in Shapiro–Wilk test

CV coefficient of variation, *NS* not significant

** $P < 0.01$

at 35°40′N, 139°46′E), from 2007 to 2011 excluding 2009. OA-RIL population was grown at Akita in 2007, 2008, and 2010, and at Kyoto only in 2010, while ST-RIL population was grown at Akita in 2010 and 2011, and at Kyoto only in 2010 (Table 1). The experimental field at Akita has a highly humic andosol and was fertilized with 24 kg ha^{-1} of N, 80 kg ha^{-1} of P, and 80 kg ha^{-1} of K before sowing; each line or cultivar was planted in plots 2–2.4 m in length with a row spacing of 0.75 m and plant separation within each row of 0.12 m. The experimental field at Kyoto has an alluvial sandy loam soil and was fertilized with 30 kg ha^{-1} of N, 100 kg ha^{-1} of P, and 100 kg ha^{-1} of K before sowing; each line or cultivar was planted in a single row plot 1.35 m in length with a row spacing of 0.7 m and plant separation within each row of 0.15 m. One hundred and eighteen lines were randomly selected from each population and grown at both Akita and Kyoto in 2010, while all lines were grown in the other years. Plots at each location were arranged according to a randomized complete block design with two replications. All plants with the exception of two at both ends of each row were harvested and threshed in bulk. Experimental conditions for each year and location are summarized in Supplementary Table 1. Segregating plants of selected OA-RHLs were grown at Akita in 2010 under the same conditions as other lines, and all the plants were individually harvested.

Phenotypic measurements

After threshing in each plot with a thresher, 200 healthy seeds were randomly selected and air-dried. Each air-dried sample was weighed, and its water content was determined with a grain moisture tester (PM830-2; Kett Electric Laboratory, Tokyo, Japan). The mean SSW was adjusted to reflect a water content of 15 %. For OA-RHLs, SSW was determined by weighing 50 randomly selected healthy seeds without adjustment for moisture content, given that the seed mass of most plants was insufficient for its measurement.

In addition to SSW, seed shape characters were determined for the OA-RIL population grown at Akita in 2007 and 2008. Ten seeds were sampled randomly from each plot for the measurement of seed height (SH), seed length (SL), and seed thickness (ST) with a digital caliper (CD-15PSX; Mitutoyo, Kanagawa, Japan). SH was defined as the longest distance vertical to the hilum, SL as the longest distance perpendicular to the SH axis, and ST as the longest distance perpendicular to the SH and SL measurements (Nelson and Wang 1989). The height to thickness ratio (HTR) was calculated by dividing SH by ST, and the height to length ratio (HLR) was calculated by dividing SH by SL (Nelson and Wang 1989). Seed volume was estimated as SL × SH × ST (Salas et al. 2006).

Agricultural characters such as flowering time (R1), the pod number per plant (PN) and seed number per pod (SNP), were evaluated in addition to SSW for the two RIL populations grown in 2010. The date on which more than 50 % of the plants flowered was recorded as the flowering date. This benchmark corresponded to the reproductive stage 1 (R1) described by Fehr et al. (1971). Pods having seeds of five plants in each plots were counted to calculate PN. The number of potential seeds for each plant was obtained in accordance with the previous report of Jeong et al. (2012) to calculate SNP. The seed number per plant (SN) was calculated by multiplying SNP and PN. Similarly, the seed weight per plant was calculated by multiplying SSW and SN.

Construction of genetic linkage maps

Total DNA was isolated from seeds of the F_6 generation of the OA-RIL population and from those of the F_7 generation of the ST-RIL population with the use of an Automatic DNA Isolation System according to Plant DNA Extraction

Protocol version 2 (PI-50a; Kurabo, Osaka, Japan). A whole-genome simple sequence repeat (SSR, or microsatellite) panel consisting of 304 such loci was applied to mapping in each population (Sayama et al. 2011). Three additional markers (GMES4177, Sat_092, Sat_222) were selected from the integrated linkage map (Hwang et al. 2009), and a phenotypic marker (*Dt1*) of stem growth habit was also used to construct the linkage maps for both the OA-RIL and ST-RIL populations. In addition, the genotype of a leaf shape locus (*ln*) in each line of the ST-RIL population was analyzed on the basis of a cleaved amplified polymorphic sequence (CAPS) marker with the use of a pair of polymerase chain reaction (PCR) primers (5′-CCCTT TAACTTTCTCTCTCTTATGAC-3′, 5′-GGTATGATCAT GAAGTTTACCGGA-3′) and the restriction enzyme *Bgl*II. The *Ln* genotype harbors an AGATCT sequence recognized by *Bgl*II in the region of the base sequence determining leaf shape, whereas the *ln* genotype harbors an ACATCT sequence at the same position (Jeong et al. 2012). The PCR products were reacted with 1 U of *Bgl*II (Toyobo, Tokyo, Japan) for 5 h at 37 °C and then fractionated as described previously (Hwang et al. 2008). Linkage between markers was analyzed with the use of MAPMAKER/EXP 3.0b (Lincoln 1992). Genetic distances (centimorgans) were calculated with the Kosambi mapping function (Kosambi 1943). Linkage maps were graphically visualized with MapChart (Voorrips 2002).

QTL analysis

The position and effects of QTLs were calculated by composite interval mapping (CIM) analysis with the use of Windows QTL Cartographer version 2.5 (Wang et al. 2010) and with Model 6 (Standard Model and Regression method 2; Backward Regression method with the default parameters). The genome was scanned at 1 cM intervals. For CIM analysis, we conducted 1,000 × permutation tests (significance level, $P < 0.05$) to determine the LOD (logarithm of odds) score threshold for the detection of minor QTL effects beyond the effects of large environmental differences. Detected QTLs including the same adjacent markers were regarded as identical in this study.

Statistical analysis

Quantitative data are presented as mean ± SD. Statistical analysis was performed with the statistical package R version 2.12.2 (http://www.r-project.org). The statistical significance of differences was evaluated by analysis of variance (ANOVA) followed by Student's *t* test for comparisons between two groups or by the Tukey–Kramer test for comparisons among three or more groups. Pairwise associations were examined by Pearson product-moment correlation analysis. A *P* value of <0.05 was considered statistically significant. The normality of phenotypic data was evaluated by the Shapiro–Wilk test.

Results

Phenotypic variation of mean seed weight

As expected, SSW differed between parental cultivars, with that for Ohsuzu and Tachinagaha being about twice that for Athow and Stressland, respectively, in all environments (Table 1). SSW in the two RIL populations showed a continuous distribution that approximately spanned the values for the two parents (Supplementary Fig. 1). In the OA-RIL population, 29 lines in 2007, 2 lines in 2008, 5 lines in 2010 at Akita, and 8 lines in 2010 at Kyoto showed abnormal growth as a result of physiological disorders or disease and were therefore excluded from subsequent analysis, whereas 18 lines in 2010 at Akita, 12 lines in 2010 at Kyoto, and 57 lines in 2011 were excluded for the same reasons in the ST-RIL population. There was a significant positive correlation ($P < 0.001$) for SSW in each combination of environments for each population (Supplementary Fig. 2, Supplementary Table 2). SSW for the all data sets except the data set of ST-RILs in 2010 at Kyoto exhibited a normal distribution (Table 1). ANOVA for comparison of SSW in two different pairs of data sets for the OA-RIL population and between two datasets for the ST-RIL population revealed significant effects for RIL (G) and cultivation year (E) as well as a significant interaction between RIL and cultivation year (G × E) at the $P < 0.01$ or $P < 0.001$ level (Supplementary Tables 3–5). These results suggested that SSW is a typical quantitative trait after taking into account the environmental differences.

Construction of genetic linkage maps

Genetic linkage maps were constructed for each RIL population with the use of a whole-genome panel consisting of 304 SSR loci as well as three additional markers (GMES4177, Sat_092, Sat_222) (Hwang et al. 2009; Sayama et al. 2011). A total of 189 marker loci, including 185 SSRs from the whole-genome panel, the three additional markers (GMES4177, Sat_092, Sat_222), and a phenotypic marker (*Dt1*) of stem growth habit, exhibited obvious polymorphism in the OA-RIL population. The resultant linkage map for the OA-RIL population comprised 20 linkage groups (LGs) and covered a total of 2,609.1 cM. The average number of marker loci in each LG was 9.5, and the average distance between markers was 13.8 cM. For the ST-RIL population, a total of 170 marker loci, including 167 markers from the whole-genome SSR panel, one of the

Table 2 Detection of QTLs associated with single seed weight (SSW) in the OA-RIL population grown in four different environments

Cultivation year	Location	Chr (LG)	Position (cM)	Marker interval			LOD score[a]	Effect[b] (mg)	R2[c] (%)	Name of QTL[d]
2007	Akita	17 (D2)	45.5	Satt372	–	GMES4177	8.9	14.9	16.0	***qSw17-1*$_{OA}$**
		4 (C1)	95.1	Sat_235	–	AI794821	3.2	−8.4	5.2	
2008	Akita	17 (D2)	45.5	Satt372	–	GMES4177	14.1	14.3	20.9	***qSw17-1*$_{OA}$**
		18 (G)	45.0	Sat_315	–	Satt394	3.6	8.2	7.0	
		9 (K)	57.1	Satt559	–	Satt499	3.6	6.8	4.8	
		15 (E)	75.4	Satt263	–	Sat_380	3.6	7.4	4.8	*qSw15-1*$_{OA}$
		13 (F)	82.7	Satt663	–	Satt114	3.4	6.9	4.8	*qSw13-1*$_{OA}$
2010	Akita	17 (D2)	45.8	GMES4177	–	CSSR172	7.6	12.6	20.4	***qSw17-1*$_{OA}$**
	Kyoto	15 (E)	83.8	Satt263	–	Sat_380	5.2	11.2	13.2	*qSw15-1*$_{OA}$
		4 (C1)	51.5	GMES0780	–	Satt646	4.8	−10.9	12.5	
		10 (O)	39.7	Satt653	–	Satt345	4.1	11.7	14.1	
		17 (D2)	45.5	Satt372	–	GMES4177	3.8	9.2	9.4	***qSw17-1*$_{OA}$**
		2 (D1b)	121.8	Sat_183	–	Satg001	3.3	–8.5	7.3	

[a] LOD scores were calculated by CIM. LOD score thresholds were 3.1 for Akita-2007, 3.1 for Akita-2008, 3.1 for Akita-2010, and 3.2 for Kyoto-2010, which were equivalent to a 5 % genome-wide Type 1 error rate in each environment, and they were applied to QTL detection

[b] Additive effect: a positive value indicates that the effect of the maternal (Ohsuzu) allele is to increase SSW

[c] Proportion of total phenotypic variation explained by the locus

[d] Bold type indicates a QTL detected at a similar position in all four environments

three additional markers (GMES4177), a molecular marker of leaf shape (*ln*), and the phenotypic marker *Dt1*, exhibited obvious polymorphism. The resultant linkage map for the ST-RIL population comprised 20 LGs and covered a total of 2,671.1 cM. The average number of marker loci in each LG was 8.5, and the average distance between markers was 15.7 cM. The marker order and distance between loci for the constructed linkage maps accorded well with those of our previous maps (Hwang et al. 2009; Sayama et al. 2011). We therefore applied the generated linkage maps to the detection of QTLs for SSW.

QTL analysis for SSW in different environments

Proper LOD thresholds were calculated on the basis of a genome-wide permutation test (1000×) for the identification of significant QTLs for SSW in both OA-RIL and ST-RIL populations. Nine QTLs for SSW with LOD scores ranging 3.2–14.1 were mapped to chromosomes 2, 4, 9, 10, 13, 15, 17, and 18 in the OA-RIL population, with each of these loci accounting for 4.8–20.9 % of total phenotypic variation (Table 2). A QTL that mapped to chromosome 17 (adjacent to GMES4177) was consistent across environments and was designated *qSw17-1*$_{OA}$ (quantitative trait locus for single seed weight on chromosome 17–number 1 in the OA-RIL population) (Fig. 1a). The LOD scores for this locus (8.9, 14.1, 7.6, and 3.8) were the largest of those obtained in three of the four environments, and the locus accounted for 16.0, 20.9, 20.4, and 9.4 % of phenotypic variation in the respective environments. Another QTL that mapped between Satt263 and Sat_380 on chromosome 15 were detected in two environments, and was designated as *qSw15-1*$_{OA}$. For Kyoto-2010, *qSw15-1*$_{OA}$ yielded the largest LOD score (5.2) and accounted for 13.2 % of phenotypic variation. These two QTLs, *qSw17-1*$_{OA}$ and *qSw15-1*$_{OA}$, had positive additive effects on SSW, indicating that Ohsuzu contributed alleles that led to an increase in SSW. The remaining seven QTLs were detected only one of the four environments in the OA-RIL population.

For the ST-RIL population, six QTLs for SSW with LOD scores of 3.5–18.5 were mapped to chromosomes 5, 12, 13, 17, and 20, with each locus accounting for 4.3–27.0 % of total phenotypic variation (Table 3). A QTL that mapped to the same location as *qSw17-1*$_{OA}$, designated *qSw17-1*$_{ST}$, was consistent across environments (Fig. 1b). These findings suggested that *qSw17-1*$_{ST}$ and *qSw17-1*$_{OA}$ might be identical and were therefore together designated *qSw17-1*. The *qSw17-1*$_{ST}$ locus had the second highest LOD score (3.6, 4.8, and 11.8) in each of the three environments, and it accounted for 11.4, 11.6, and 15.0 % of phenotypic variation in these respective environments. The Japanese cultivar Tachinagaha contributed to the increase in SSW attributable to *qSw17-1*$_{ST}$. Another significant QTL, which mapped in the vicinity of Sat_105 on chromosome 20 and was designated *qSw20-1*$_{ST}$, was consistent across environments and yielded the highest LOD score (7.1, 7.6, and 18.5) in each, and it accounted for 26.0, 20.3, and 27.0 % of phenotypic variation in the respective environments. In contrast to *qSw17-1*$_{ST}$, the US cultivar Stressland provided alleles of *qSw20-1*$_{ST}$ responsible for the increase

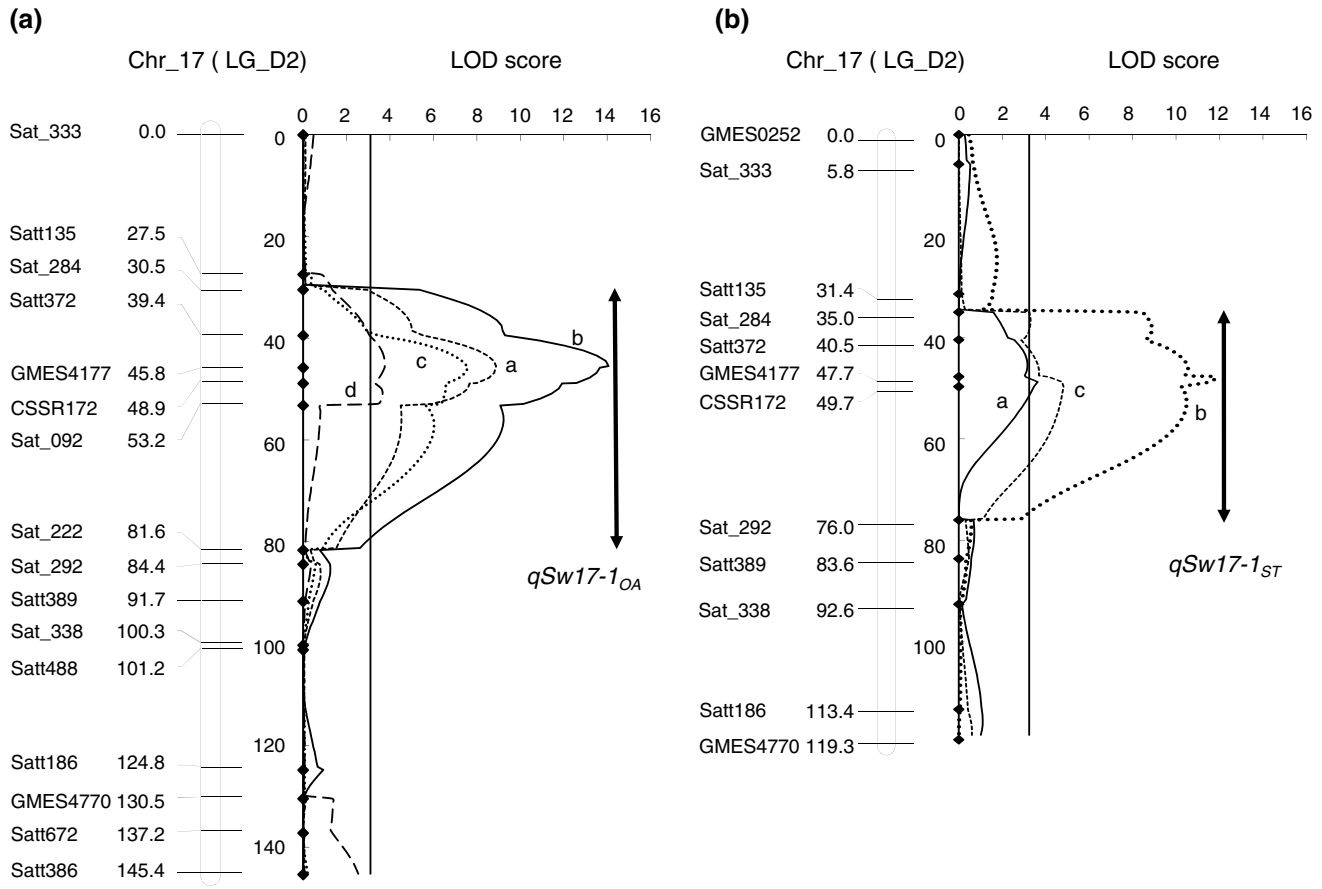

Fig. 1 Linkage maps of chromosome 17 (LG D2) for a QTL associated with single seed weight (SSW) (*qSw17-1*) in the OA-RIL and ST-RIL populations. **a** LOD scores for the QTL in the OA-RIL population grown in four different environments (Akita-2007, Akita-2008, Akita-2010, and Kyoto-2010) are indicated with the *lines* labeled *a*, *b*, *c*, and *d*, respectively. The LOD threshold (3.1) for Akita-2007 is also indicated by a continuous line. The double-headed arrow denotes the location of $qSw17\text{-}1_{OA}$. Marker names together with the cumulative map distance from Sat_333 (in centimorgans) are shown. **b** LOD scores for the QTL in the ST-RIL population grown in three different environments (Akita-2010, Akita-2011, and Kyoto-2010) are indicated with the *lines* labeled *a*, *b*, and *c*, respectively. The LOD threshold (3.2) for Akita-2010 is also indicated by a continuous line. The double-headed arrow denotes the location of $qSw17\text{-}1_{ST}$. Marker names together with the cumulative map distance from the marker GMES0252 (in centimorgans) are shown

Table 3 Detection of QTLs associated with single seed weight (SSW) in the ST-RIL population grown in three different environments

Cultivation year	Location	Chr (LG)	Position (cM)	Marker interval			LOD score[a]	Effect[b] (mg)	R2[c] (%)	QTL[d]
2010	Akita	20 (I)	59.3	Sat_105	–	*Ln*	7.1	16.7	26.0	***qSw20-1*$_{ST}$**
		17(D2)	48.8	GMES4177	–	CSSR172	3.6	−11.1	11.4	***qSw17-1*$_{ST}$**
	Kyoto	20 (I)	58.3	Sat_105	–	*Ln*	7.6	15.8	20.3	***qSw20-1*$_{ST}$**
		17(D2)	49.7	GMES4177	–	CSSR172	4.8	−12.3	11.6	***qSw17-1*$_{ST}$**
		13 (F)	82.7	Sat_039	–	Satt663	3.6	−10.3	8.6	*qSw13-1*$_{ST}$
2011	Akita	20 (I)	60.3	Sat_105	–	*Ln*	18.5	17.8	27.0	***qSw20-1*$_{ST}$**
		17(D2)	47.8	GMES4177	–	CSSR172	11.8	−12.9	15.0	***qSw17-1*$_{ST}$**
		12 (H)	73.6	Sat_206	–	Satt302	5.7	−9.6	8.4	
		20 (I)	100.7	Satt623	–	Sct_189	4.7	−8.0	5.6	
		5 (A1)	18.0	Sat_368	–	Sat_344	3.5	−6.9	4.3	

[a] LOD score thresholds of 3.2 for Akita-2010, 3.2 for Kyoto-2010, and 3.2 for Akita-2011 were applied to QTL detection

[b] Additive effect: a positive value indicates that the maternal (Stressland) allele increases

[c] Proportion of total phenotypic variation explained by the locus

[d] Bold type indicates QTLs detected at a similar position in all three environments

in SSW. A genic marker (narrow leaflet, *ln*) derived from Tachinagaha mapped at the same position as $qSw20\text{-}1_{ST}$ (Supplementary Fig. 3).

A QTL, $qSw13\text{-}1_{OA}$, was detected between Satt663 and Satt114 on chromosome 13 in the OA-RIL population for Akita-2008, and a QTL, $qSw13\text{-}1_{ST}$, was observed in a similar region in the ST-RIL population for Kyoto-2010 (Tables 2 and 3). The Japanese cultivars provided the alleles of both loci responsible for the increase in SSW. These findings suggested that $qSw13\text{-}1_{OA}$ and $qSw13\text{-}1_{ST}$ might be identical and were therefore together tentatively designated *qSw13-1*. In contrast to *qSw17-1*, however, this QTL was unstable, being detected only once in each recombinant inbred population, in which it exhibited LOD scores of 3.4 and 3.6 and accounted for 4.8 and 8.6 % of phenotypic variation, respectively. The remaining three QTLs were detected in only one of the three environments in the ST-RIL population.

Thus, among the four QTLs (*qSw17-1*, *qSw20-1*, *qSw15-1*, and *qSw13-1*), only *qSw17-1* appeared to be stable across all environments and genetic backgrounds (Tables 2 and 3).

Mode of inheritance of *qSw17-1*

To investigate the mode of inheritance of *qSw17-1*, we selected two RHLs from the OA-RIL population according to a heterozygous region adjacent to this QTL. The progeny of the RHLs would be expected to show simple phenotypic segregation based on the effects of the target QTL at the residual heterozygous region in a uniform genetic background (Yamanaka et al. 2005). The progeny of the two OA-RHLs was thus divided into three groups (Ohsuzu type, Athow type, and heterozygous type) according to the genotype of the SSR marker GMES4177, which is proximal to *qSw17-1*. The Ohsuzu-type and heterozygous-type progeny showed a significantly increased SSW compared with the Athow-type progeny (Table 4). The genotype of *qSw17-1* was thus found to contribute significantly to the variation in SSW, with a dominant allele derived from Ohsuzu being responsible for the increase in SSW.

Effects of *qSw17-1* genotype on seed shape

Soybean has a highly variable seed shape as well as SSW. To investigate the possible effects of *qSw17-1* genotype on seed shape, we determined characters related to seed shape in the OA-RIL population grown at Akita in 2007 and 2008 (Table 5). For all characters examined, the effect of line (G), the effect of location (E), and the interaction between line and location (G × E) were significant at the 0.1 or 1 % level (Supplementary Tables 3), suggesting that these characters are subject to genetic control beyond environmental differences. The individual lines of the OA-RIL population were divided into two allelic types, Ohsuzu type and Athow type, for *qSw17-1* on the basis of the genotype of GMES4177, and the seed shape characters were compared between the two allelic types. The genotype of GMES4177 contributed significantly to the variation in SSW, with Ohsuzu contributing a positive allele. Similarly, seed volume (SV), seed height (SH), seed length (SL), and seed thickness (ST) differed significantly between the two *qSw17-1* allelic types, but the corresponding height to thickness (HTR) and height to length (HLR) ratios did not. These results, thus, indicated that *qSw17-1* contributes to a uniform enlargement of soybean seed with no effect on seed shape per se.

Table 4 Comparison of single seed weight (SSW) for two RHLs segregating *qSw17-1* genotype

Parents or RHLs selected from OA-RILs	Genotype at *qSw17-1* (GMES4177)[a]	No. of plants	SSW (mg/seed)[b] Mean ± SD
P1: Ohsuzu		10	282 ± 24
P2: Athow		10	186 ± 9
OA-RIL-217	Ohsuzu	51	210 ± 18A
	Heterozygote	110	207 ± 19A
	Athow	62	195 ± 22B
OA-RIL-221	Ohsuzu	8	256 ± 20A
	Heterozygote	14	249 ± 9A
	Athow	13	227 ± 62B

[a] Allelic type for *qSw17-1* is based on the genotype of the marker GMES4177

[b] Values followed by the same uppercase letter do not differ significantly at the 5 % level from each other within each RHL (Tukey–Kramer multiple-comparison test)

Effects of *qSw17-1* genotype on agricultural characters

We also examined other agricultural characters such as SNP, PN and flowering time (R1) for the OA-RIL and ST-RIL populations grown at the two locations (Akita and Kyoto) in 2010. For all characters examined, the effect of line (G), the effect of location (E), and the interaction between line and location (G × E) were significant at the 0.1 or 1 % level (Supplementary Tables 4 and 5), suggesting that these characters are subject to genetic control beyond environmental differences. The individual lines of the RIL populations were also divided into two parental allelic types for *qSw17-1* on the basis of the genotype of the adjacent marker GMES4177, and the agricultural characters were then compared between the two allelic types (Supplementary Tables 6 and 7). The genotype of *qSw17-1* was found to contribute significantly to SSW, with Ohsuzu and Tachinagaha contributing positive alleles to the OA-RIL population at both locations and to the ST-RIL

Table 5 Relation between characters related to seed shape and *qSw17-1* genotype for the OA-RIL population grown at Akita in 2007 and 2008

Character	Akita-2007					Akita-2008				
	P1 Ohsuzu	P2 Athow	Ohsuzu type[a] ($n = 95$) Mean ± SD	P[b]	Athow type[a] ($n = 96$) Mean ± SD	P1 Ohsuzu	P2 Athow	Ohsuzu type[a] ($n = 108$) Mean ± SD	P[b]	Athow type[a] ($n = 110$) Mean ± SD
Single seed weight (mg)	373	200	290 ± 36	***	263 ± 32	302	177	247 ± 30	***	221 ± 26
Seed height (mm)	8.89	7.37	8.49 ± 0.38	***	8.27 ± 0.35	8.43	7.06	8.02 ± 0.42	***	7.75 ± 0.40
Seed length (mm)	8.38	6.87	7.91 ± 0.28	***	7.66 ± 0.27	7.84	6.59	7.58 ± 0.38	***	7.26 ± 0.33
Seed thickness (mm)	7.19	5.68	6.74 ± 0.28	***	6.53 ± 0.29	6.80	5.46	6.26 ± 0.34	***	6.02 ± 0.35
HTR[c]	1.17	1.21	1.18 ± 0.04	NS	1.18 ± 0.04	1.15	1.21	1.21 ± 0.04	NS	1.21 ± 0.04
HLR[d]	0.94	0.93	0.93 ± 0.03	NS	0.93 ± 0.03	0.93	0.93	0.95 ± 0.03	NS	0.94 ± 0.03
Seed volume (mm^3)[e]	536	288	454 ± 50	***	415 ± 43	449	254	383 ± 54	***	340 ± 47

[a] Allelic type for *qSw17-1* is based on the genotype of the marker GMES4177

[b] Significance of differences between the Ohsuzu allele and the Athow allele: *** $P < 0.001$ (Student's t test); *NS* not significant

[c] Height to thickness ratio (HTR)

[d] Height to length ratio (HLR)

[e] Seed volume was calculated by multiplying length, height, and thickness

population at Kyoto, respectively. Other characters including R1 did not differ significantly between the *qSw17-1* genotypes in any of the four data sets.

Discussion

We have performed a detailed analysis of SSW and genotypes in two RIL populations of soybean derived from parents that manifest a large difference in SSW. We identified four QTLs (*qSw17-1*, *qSw20-1*, *qSw15-1*, and *qSw13-1*) that were associated with SSW consistently across environments or genetic backgrounds. Among these QTLs, only *qSw17-1* was identified in all environments for both populations, and it accounted for a large proportion (9.4–20.9 %) of total phenotypic variation in SSW. To date, many QTLs for SSW that map to all 20 chromosomes have been described in SoyBase. Some of these previously described QTLs were mapped to chromosome 17 (LG D2), but they were identified with a lower LOD score compared with that for *qSw17-1* (Hoeck et al. 2003; Hyten et al. 2004; Panthee et al. 2005; Liu et al. 2007; Han et al. 2012). Fine-scale mapping will be required to determine whether any of these previously identified loci are identical to *qSw17-1*. Most of the populations analyzed for SSW in previous studies were developed from parents with relatively small differences in this trait. In contrast, we developed segregating populations from US and Japanese cultivars that differ in SSW by a factor of ~2. Japanese soybean germplasm was previously found to show wide variation in SSW (Kaga et al. 2012), given that seed size is a determinative factor for soy food products such as natto, boiled beans, and edamame. Indeed, the *qSw17-1* alleles of the large-seed parental cultivars, Ohsuzu and Tachinagaha, contributed to increasing SSW in the RIL populations of the present study. Some Japanese soybean accessions have a SSW (>700 mg per seed) much larger than those of the two Japanese cultivars used in our study. Japanese germplasm thus likely harbors multiple alleles and loci that contribute to the variation in SSW.

Soybean cultivars also show diversity in seed shape (Nelson and Wang 1989; Kaga et al. 2012), which is also under genetic control (Cober et al. 1997). Seed shape is also associated with seed quality for food-type soybean, given that round seeds are mainly used as a food source. We found that, although *qSw17-1* was a determinant of seed size, it did not affect seed shape, consistent with a previous observation that seed shape and seed size are not correlated (Cober et al. 1997).

SSW is an important factor in seed yield, with positive relations between seed size and seed yield having been described for soybean (Burris et al. 1973; Smith and Camper 1975). However, we did not detect any significant QTL associated with seed yield in the vicinity of *qSw17-1* (data not shown). We divided the individual lines of the OA-RIL and ST-RIL populations into two groups according to *qSw17-1* genotype as estimated from the genotype of GMES4177 (Supplementary Tables 6 and 7). The genotypes of *qSw17-1* manifested a significant difference in SSW, indicating that GMES4177 might be a powerful molecular marker for the selection of SSW. The *qSw17-1* genotypes did not manifest a significant difference in total seed weight, however. Changes in other yield components such as pod number per plant and seed number per plant appeared to offset the increase in SSW, although no significant differences in these parameters or in seed number per pod were detected between the two genotypes. A similar phenomenon was

previously observed for two QTLs associated with a yield component in rice (Ohsumi et al. 2011). The introduction of two QTLs that increase spikelet number per panicle, thus, had a little impact on grain yield as a result of compensatory changes in other yield components. An increase in sink size such as SSW, thus, does not necessarily result in a substantial improvement in seed yield (Sadras 2007). Fundamental traits such as photosynthetic ability for provision of a sufficient source supply would thus likely need to be improved to achieve a response to an increase in sink demand.

In addition to *qSw17-1*, we identified three other QTLs (*qSw20-1*, *qSw15-1*, and *qSw13-1*) that impacted SSW in at least two environments and or both populations. Tachinagaha has a narrow leaflet phenotype controlled by a single gene, *ln* (Bernard and Weiss 1973), whereas Stressland has an ovate (wild) leaflet phenotype controlled by a dominant allele (*Ln*) of this gene. A genic marker for the *ln* locus mapped to the same position as *qSw20-1* in the present study. The narrow leaflet allele in the homozygous condition also increases SNP (Hartwig and Edwards 1970; Mandl and Buss 1981; Jeong et al. 2012), and this allele has a negative effect on SSW as a result of the increase in SNP (Dinkins et al. 2002). Stressland, which harbors the *Ln* allele, provided alleles of *qSw20-1* that led to an increase in SSW and a decrease in SNP (Supplementary Tables 8). It thus appears likely that *qSw20-1* is identical to *ln*. In contrast, *qSw15-1* and *qSw13-1*, like *qSw17-1*, did not appear to affect leaflet phenotype or SNP (data not shown), suggesting that the functions of the responsible genes differ from that of *ln*. The *qSw15-1* locus was associated with SSW in the OA-RIL population in two environments, with the Ohsuzu allele increasing seed weight. QTLs for SSW were previously detected on chromosome 15 (Mian et al. 1996; Orf et al. 1999), but their relation to *qSw15-1* remains unclear. The QTL *qSw13-1* was detected only once in each recombinant inbred population, and, as with *qSw15-1* and *qSw17-1*, the alleles of this QTL derived from the Japanese cultivars increased SSW. QTLs were also previously identified in the vicinity of *qSw13-1* (Mian et al. 1996; Hoeck et al. 2003; Hyten et al. 2004). Unlike *qSw17-1* and *qSw20-1*, however, *qSw15-1* and *qSw13-1* appeared to be unstable and greatly influenced by environment.

Our two RIL populations allowed the detection of several QTLs associated with SSW in soybean, among which *qSw17-1* was found to be stable and to have a substantial effect across populations and environments. Our results reflect the great variation within and large differences between US and Japanese cultivars with regard to SSW (Kaga et al. 2012). Further characterization of *qSw17-1*, including isolation of the responsible gene, should provide further insight into such differences in SSW as well as provide a basis for the development of breeding lines with favorable alleles by marker-assisted selection.

Acknowledgments This study was supported by grants from the Ministry of Agriculture, Forestry, and Fisheries of Japan [Genomics for Agricultural Innovation (DD-3260, SOY2002, and SOY2004) and Genomics-Based Technology for Agricultural Improvement (SFC-1001)]. We thank Yumi Nakamoto of National Agricultural Research Center for Hokkaido Region and Chiaki Kato of National Agricultural Research Center for Tohoku Region for their assistance with data processing for our QTL analysis.

Conflict of interest The authors declare that they have no conflict of interest.

References

Abe J, Xu DH, Suzuki Y, Kanazawa A, Shimamoto Y (2003) Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor Appl Genet 106:445–453

Bernard R, Weiss M (1973) Qualitative genetics. In: Caldwell BE (ed) Soybeans: improvement, production, and uses. American Society of Agronomy, Madison, pp 117–154

Brim CA (1973) Quantitative genetics and breeding. In: Caldwell BE (ed) Soybeans: improvement, production, and uses. American Society of Agronomy, Madison, pp 155–186

Brim CA, Cockerham CC (1961) Inheritance of quantitative characters in soybeans. Crop Sci 1:187–190

Burris JS, Edje OT, Wahab AH (1973) Effects of seed size on seedling performance in soybeans: II. Seedling growth and photosynthesis and field performance. Crop Sci 13:207–210

Chapman A, Pantalone VR, Ustun A, Allen FL, Landau-Ellis D, Trigiano RN, Gresshoff PM (2003) Quantitative trait loci for agronomic and seed quality traits in an F_2 and $F_{4:6}$ soybean population. Euphytica 129:387–393

Cober ER, Voldeng HD, Fregeau-Reid JA (1997) Heritability of seed shape and seed size in soybean. Crop Sci 37:1767–1769

Cooper RL, Martin RJ, Martin SK, Calip-DuBois A, Fioritto RJ, Schmitthenner AF (1999) Registration of 'Stressland' soybean. Crop Sci 39:590–591

Dinkins RD, Keim KR, Farno L, Edwards LH (2002) Expression of the narrow leaflet gene for yield and agronomic traits in soybean. J Heredity 93:346–351

Edwards CJ, Hartwig EE (1971) Effect of seed size upon rate of germination in soybeans. Agron J 63:429–450

Fasoula VA, Harris DK, Boerma HR (2004) Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. Crop Sci 44:1218–1225

Fehr W, Caviness C, Burmood D, Pennington J (1971) Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. Crop Sci 11:929–931

Han Y, Li D, Zhu D, Li H, Li X, Teng W, Li W (2012) QTL analysis of soybean seed weight across multi-genetic backgrounds and environments. Theor Appl Genet 125:671–683

Hartwig EE (1973) Varietal development. In: Caldwell BE (ed) Soybeans: improvement, production, and uses. American Society of Agronomy, Madison, pp 187–210

Hartwig EE, Edwards CJ (1970) Effects of morphological characteristics upon seed yield in soybeans. Agron J 62:64–65

Hoeck JA, Fehr WR, Shoemaker RC, Welke GA, Johnson SL, Cianzio SR (2003) Molecular marker analysis of seed size in soybean. Crop Sci 43:68–74

Hopper NW, Overholt JR, Martin JR (1979) Effect of cultivar, temperature and seed size on the germination and emergence of soya beans (*Glycine max* (L.) Merr.). Ann Bot 44:301–308

Hu Z, Zhang H, Kan G, Ma D, Zhang D, Shi G, Hong D, Zhang G, Yu D (2013) Determination of the genetic architecture of seed size

and shape via linkage and association analysis in soybean (*Glycine max* L. Merr.). Genetica 141:247–254

Hwang TY, Nakamoto Y, Kono I, Enoki H, Funatsuki H, Kitamura K, Ishimoto M (2008) Genetic diversity of cultivated and wild soybeans including Japanese elite cultivars as revealed by length polymorphism of SSR markers. Breed Sci 58:315–323

Hwang TY, Sayama T, Takahashi M, Takada Y, Nakamoto Y, Funatsuki H, Hisano H, Sasamoto S, Sato S, Tabata S, Kono I, Hoshi M, Hanawa M, Yano C, Xia Z, Harada K, Kitamura K, Ishimoto M (2009) High-density integrated linkage map based on SSR markers in soybean. DNA Res 16:213–225

Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, Schmidt ME (2004) Seed quality QTL in a prominent soybean population. Theor Appl Genet 109:552–561

Hyten DL, Choi I-Y, Song Q, Specht JE, Carter TE, Shoemaker RC, Hwang E-Y, Matukumalli LK, Cregan PB (2010) A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. Crop Sci 50:960–968

Jeong N, Suh SJ, Kim M-H, Lee S, Moon J-K, Kim HS, Jeong S-C (2012) *Ln* is a key regulator of leaflet shape and number of seeds per pod in soybean. Plant Cell 24:4807–4818

Kaga A, Shimizu T, Watanabe S, Tsubokura Y, Katayose Y, Harada K, Vaughan DA, Tomooka N (2012) Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. Breed Sci 61:566–592

Kosambi D (1943) The estimation of map distances from recombination values. Ann Eugenics 12:172–175

Lincoln S (1992) Constructing genetic maps with MAPMAKER/EXP 3.0. Tecnical report of the Whitehead Institute, Cambridge

Liu B, Fujita T, Yan Z-H, Sakamoto S, Xu D, Abe J (2007) QTL mapping of domestication-related traits in soybean (*Glycine max*). Ann Bot 100:1027–1038

Mandl F, Buss G (1981) Comparison of narrow and broad leaflet isolines of soybean. Crop Sci 21:25–27

Mian MAR, Bailey MA, Tamulonis JP, Shipe ER, Carter TE, Parrott WA, Ashley DA, Hussey RS, Boerma HR (1996) Molecular markers associated with seed weight in two soybean populations. Theor Appl Genet 93:1011–1016

Nelson RL, Wang P (1989) Variation and evaluation of seed shape in soybean. Crop Sci 29:147–150

Niu Y, Xu Y, Liu X-F, Yang S-X, Wei S-P, Xie F-T, Zhang Y-M (2013) Association mapping for seed size and shape traits in soybean cultivars. Mol Breed 31:785–794

Ohsumi A, Takai T, Ida M, Yamamoto T, Arai-Sanoh Y, Yano M, Ando T, Kondo M (2011) Evaluation of yield performance in rice near-isogenic lines with increased spikelet number. Field Crops Res 120:68–75

Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, Lark KG (1999) Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. Crop Sci 39:1642–1651

Panthee DR, Pantalone VR, West DR, Saxton AM, Sams CE (2005) Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. Crop Sci 45:2015–2022

Sadras VO (2007) Evolutionary aspects of the trade-off between seed size and number in crops. Field Crops Res 100:125–138

Salas P, Oyarzo-Llaipen JC, Wang D, Chase K, Mansur L (2006) Genetic mapping of seed shape in three populations of recombinant inbred lines of soybean (*Glycine max* L. Merr.). Theor Appl Genet 113:1459–1466

Sayama T, Hwang T-Y, Komatsu K, Takada Y, Takahashi M, Kato S, Sasama H, Higashi A, Nakamoto Y, Funatsuki H, Ishimoto M (2011) Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean. DNA Res 18:107–115

Smith T, Camper H (1975) Effects of seed size on soybean performance. Agron J 67:681–684

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Ude GN, Kenworthy WJ, Costa JM, Cregan PB, Alvernaz J (2003) Genetic diversity of soybean cultivars from China, Japan, North America, and North American ancestral lines determined by amplified fragment length polymorphism. Crop Sci 43:1858–1867

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Heredity 93:77–78

Wang S, Basten CJ, Zeng ZB (2010) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh

Watanabe S, Tajuddin T, Yamanaka N, Hayashi M, Harada K (2004) Analysis of QTLs for reproductive development and seed quality traits in soybean using recombinant inbred lines. Breed Sci 54:399–407

Wilcox JR, Abney TS (1997) Registration of 'Athow' soybean. Crop Sci 37:1981–1982

Xu Y, Li H-N, Li G-J, Wang X, Cheng L-G, Zhang Y-M (2011) Mapping quantitative trait loci for seed size traits in soybean (*Glycine max* L. Merr.). Theor Appl Genet 122:581–594

Yamanaka N, Watanabe S, Toda K, Hayashi M, Fuchigami H, Takahashi R, Harada K (2005) Fine mapping of the *FT1* locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. Theor Appl Genet 110:634–639

Zhang WK, Wang YJ, Luo GZ, Zhang JS, He CY, Wu XL, Gai JY, Chen SY (2004) QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. Theor Appl Genet 108:1131–1139

Zhou X, Carter TE, Cui Z, Shoji M, Burton JW (2000) Genetic base of Japanese soybean cultivars released during 1950–1988. Crop Sci 40:1794–1802